

## 研究报告 Original Papers

## 拟南芥ABC转运类蛋白家族的分子进化、表达模式和蛋白功能网络预测分析

朱璐, 许杰, 张大兵\*

上海交通大学生命科学技术学院, 上海200240

**摘要:** ABC转运蛋白又称腺苷三磷酸结合盒转运蛋白(ATP-binding cassette transporters), 该基因家族是目前已知最大、最古老的蛋白家族之一, 在植物中ABC转运蛋白种类繁多、结构复杂、功能多样, 涉及植物一切的生命活动过程。本研究系统介绍了拟南芥中131个ABC转运蛋白的亚家族分类、系统命名、蛋白大小以及蛋白亚细胞定位等基因信息, 在此基础上, 分析了ABC转运蛋白基因在染色体分布以及进化过程中发生的复制事件; 其次在47个组织器官或发育阶段中聚类分析了ABC转运蛋白的表达模式和各个亚家族分布规律, 结果表明ABC转运蛋白基因的表达具有明显的组织特异性和时空特异性, 说明在进化过程中该类蛋白功能也进一步发生分化; 我们以花药发育过程为例, 说明ABC转运蛋白在花药发育过程中具有较高的协调性, 在时空和组织上表达受到严格的调控; 最后我们分析了ABC转运蛋白亚家族内部和各个亚家族之间可能存在的蛋白相互作用关系, 推测ABC半分子转运蛋白形成同源或异源二聚体发挥功能的可能性, 进一步说明ABC转运蛋白在蛋白互作水平上也存在功能多样性和严格的调控关系。

**关键词:** 拟南芥; ABC转运蛋白; 分子进化; 表达分析; 蛋白相互作用

**Molecular Evolution, Expression and Functional Network Prediction Analysis of ABC Transporter Gene Family in *Arabidopsis thaliana***

ZHU Lu, XU Jie, ZHANG Da-Bing\*

School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China

**Abstract:** ABC transporters (ATP-binding Cassette transporters) family is one of the biggest and oldest protein families, they have a variety of complex structure and diverse functions, participating in the various processed of plant growth and development. In this study, bioinformatics analysis identified 131 ABC transporters genes in *Arabidopsis thaliana*, a comprehensive overview of this gene family was presented, including the subfamily classification, gene names, protein size and sub-cellular localization. The 131 ABC transporters genes were grouped into 8 subfamilies on basis of phylogenetic relationship. Chromosome localization and gene duplication analysis suggest that segmental and retro-transposition-like event may result in the ABC transporters gene expansion. By analyzing the developmental expression pattern of ABC transporters genes of all the sub-families in 47 organs or development stages in *Arabidopsis*, most of the ABC transporters genes were shown to be expressed in obvious tissues, specific developmental time, suggesting that the functions of these proteins differentiated during the evolution. We used anther development as one example and expression analysis showed highly coordinated expression of the ABC transporters. Moreover we predicted the possible interactions between ABC transporter proteins in every sub-family or among different ones using bioinformatic tool and speculate the possible functional forms ABC half-transporters as homodimer or heterodimer. This work is helpful in understanding the functional role of ABC transporters in plants.

**Key words:** *Arabidopsis thaliana*; ABC transporter; molecular evolution; expression analysis; protein-protein interaction

ABC转运蛋白又称腺苷三磷酸结合盒转运蛋白(ATP-binding cassette transporters), 该蛋白家族是目前已知最大、最古老的蛋白家族之一(Verrier

收稿 2012-08-18 修定 2012-11-13  
资助 自然科学基金面上项目(31000593).  
\* 通讯作者(E-mail: zhangdb@sjtu.edu.cn; Tel: 021-34205073)。

2008)。由于这种蛋白的结构中都含有一个腺苷三磷酸(ATP)的结合盒,且具有转运物质的功能,因而被称作ABC转运蛋白。所有ABC转运蛋白的结构都含有一个高度保守的ATP结合结构域(nucleotide-binding fold, NBF)和跨膜结构域(transmembrane domains, TMDs),因此ABC转运蛋白是作为一种膜整合蛋白来行使其转运功能的。其转运的底物主要包括肽、糖、脂、重金属螯合物、多糖、生物碱、类固醇、无机离子和谷胱甘肽结合物等(Theodoulou 2000)。

ABC转运蛋白广泛存在于真核和原核生物中。目前已知人类(*Homo sapiens*)基因组中有48个ABC转运蛋白超家族成员(Borowski等2005),酵母(*Saccharomyces cerevisiae*)中有大约31个ABC转运蛋白超家族(Chen等2007)。在大肠杆菌(*Escherichia coli*) K-12基因组中,至少有80个基因编码ABC转运蛋白,约占基因组的5% (Linton和Higgins 1998)。在植物界,以水稻和拟南芥为代表的ABC转运蛋白家族的成员都多达120个以上,这是植物为适应固着生活环境,长期进化形成的(Verrier等2008)。植物中的ABC转运蛋白已经进行了统一命名,将其分为9个亚家族: ABCA~ABCI (Verrier等2008)。拟南芥基因组编码的129个ABC转运蛋白分子大小从250到1 800个氨基酸不等。根据蛋白的大小(全分子、半分子或四分之一分子)、结构与组织形式和与其他生物ABC转运蛋白的相似性比较而分为13个亚家族: MDR (multidrug resistance)、MRP/ABCC (multidrug resistance-associated protein)、PDR/ABCG (pleiotropic drug resistance)、AOH (ABC1 homolog)、PMP (peroxisomal membrane protein)、WBC/ABCG (white-brown complex homolog)、ATH (ABC2 homolog)、ATM (ABC transporter of the mitochondrion)、TAP/ABCB (transporter associated with antigen processing)、RLI (RNase L inhibitor)、GCN (general control non-repressible)、SMC (structural maintenance of chromosomes)和NAP (nonintrinsic ABC protein) (Rea 2007)。其中属于ABC全分子转运蛋白的有4个亚家族: MDR (22个成员)、MRP (15个成员)、PDR (13个成员)和AOH (1个成员);属于ABC半分子转运蛋白的有5个亚家族: PMP (2个成员)、

WBC (29个成员)、ATH (16个成员)、ATM (3个成员)和TAP (2个成员);属于ABC可溶性蛋白的有3个亚家族: RLI (2个成员)、GCN (5个成员)和SMC (4个成员)。另外还有15个可溶性蛋白没有在其他生物中发现同源蛋白而被归入NAP亚家族(Sanchez-Fernandez等2001)。

ABC蛋白作为植物细胞膜蛋白的重要组成部分,其家族成员数量多、种类广、作用方式复杂,参与植物细胞生命活动的许多过程,如植物激素运输、气孔调节、次生代谢产物的运输和环境胁迫响应以及植物与微生物互动中发挥重要作用(Rea 2007)。自从1992年国际上首次报道从拟南芥(*Arabidopsis thaliana*)中克隆的AtPGP1 (又称为AtMDR1)第一个植物ABC转运蛋白至今(Dudler和Hertig 1992),研究人员对植物ABC转运蛋白进行了多方面的研究。在植物中ABC转运蛋白的主要作用于植物次级代谢产物在特定器官中的积累,比如:根状茎(Shitan等2003)和具腺毛状体的形成(Stukkens等2005);角质层的脂类沉积(Panikashvili等2007; Pighin等2004);清除苹果酸酯和调节气孔运动(Lee等2008);解毒有害异物(Frelet-Barrand等2008; Lee等2005; Martinoia等2002)以及植物生长素的运输(Titapiwatanakun和Murphy 2009)。此外,还有研究表明,ABC转运蛋白在作物产量、品质的形成以及抗性应答中均发挥重要作用(Xu等2009)。

拟南芥作为双子叶植物分子生物学研究的模式生物,通过正向遗传学研究相应突变体的方法,是目前对于ABC转运蛋白的功能和分类研究最具有代表性和系统性的研究体系,通过对拟南芥的研究,人们对植物中ABC转运蛋白的功能有了初步的认识。但尽管如此,因为ABC转运蛋白家族在拟南芥中具有上百个成员,存在基因冗余现象(Verrier等2008),相应突变体的缺失或表型微弱难以观察、半分子蛋白互作的复杂性等等,目前完全利用正向遗传学方法,很难以对于这些基因的功能多样性进行全面系统地阐述和认识。因此,为更加全面的了解、系统的分析拟南芥中ABC转运蛋白的信息和功能,本文利用基因组测序数据与生物信息学手段,分析了拟南芥中ABC转运蛋白基因的具体信息、蛋白定位预测、染色体定位

和基因复制的进化关系,并对各个亚家族基因在不同组织器官的表达模式、在花药发育阶段的功能以及ABC转运蛋白半分子可能存在的直接相互作用进行了全面研究,旨在为拟南芥ABC转运蛋白的功能研究打下基础和提供有意义的参考价值。

## 材料与方法

### 1 拟南芥中ABC基因家族数据汇总、亚细胞定位预测和GO功能分类

从拟南芥(*Arabidopsis thaliana* L.)信息资源数据库(The *Arabidopsis* information resource, TAIR, <http://www.arabidopsis.org/>)获得拟南芥ABC超家族成员的基本信息,包括基因信息、亚家族分类、基因位点、系统学命名、别名、蛋白大小和蛋白结构域信息,进一步对比目前相关文献报道,核对基因注释的完整性和准确性。分析基因的DNA序列和蛋白结构域信息,并进一步结合在线工具ATTED-II (<http://atted.jp/>)得到蛋白的亚细胞结构预测结果。

Gene Ontology (GO)分析是使用一个基于网络的工具和数据库(Go Analysis Toolkit and Database for Agricultural community, agriGO, <http://bio-info.cau.edu.cn/agriGO/index.php>),主要包括生物过程(biological process)、分子功能(molecular function)和细胞组分(cellular component)三个部分。

### 2 ABC基因家族在染色体分布和基因进化复制分析

利用TAIR网站(<http://www.arabidopsis.org/>)的染色体图工具(chromosome map tool)作染色体图,将预测的所有ABC转运蛋白基因定位到5条染色体上的相应位置,并分析相应的基因家族连锁信息。同时利用公共资源数据库Paralogous in *Arabidopsis thaliana* (<http://oldwolfe.gen.tcd.ie/athal/dup>),分析拟南ABC转运蛋白家族成员中位于基因组倍增片段上的基因,手动添加基因之间的复制信息。

### 3 ABC转运基因表达谱信息收集和聚类分析

对获得的所有拟南芥ABC转运类蛋白基因,利用公共资源数据库(The Bio-Array Resource for Plant Biology, BAR, <http://bar.utoronto.ca/welcome>

htm)中的*Arabidopsis* eFP Browser工具,查询和下载拟南芥发育芯片表达数据,总共包括47个不同组织器官或生长发育阶段:干种子,吸胀的种子(24 h),茎的第一节点,茎叶,子叶,根,过渡期后到开花的整个莲座丛,不同花期(9、10/11、12和15期),不同时期的不同组织(花期12和花期15:雄蕊、心皮、花瓣、萼片),花梗15 d,叶1+2,7叶(叶柄、叶远侧半、叶近端半),下胚轴,根,莲座叶(2、4、6、8、10和12),衰老叶片,茎尖(花序、过渡和营养),干(第2节间),成熟花粉,种子阶段(3/4/5/6/7/8/9/10)和角果。对于与多个探针号对应的基因取其平均值作为该基因的表达值。使用微阵列数据分析Genesis (1.7.5)软件(Sturn等2002)进行分析,使用层次聚类方法对这些基因表达值进行处理,在聚类前对这些表达值以2作为底数做标准化处理。

### 4 ABC转运蛋白相互作用(PPI)分析

对获得的所有拟南芥ABC转运类蛋白基因,使用本实验室研发和已经公开发表(Wang等2012)的互动的生物信息学网络工具(An Integrated Knowledge Base for *Arabidopsis* Protein Protein Interaction Network Analyses, ANAP, [http://gmdd.shgmo.org/Computational-Biology/ANAP/ANAP\\_V1.1/](http://gmdd.shgmo.org/Computational-Biology/ANAP/ANAP_V1.1/))进行分析,分别标明各个亚家族基因的位置和关系。

## 实验结果

### 1 拟南芥ABC转运类蛋白的基本信息和GO功能分类

从拟南芥基因组注释数据库TAIR中查找并获得131个ABC家族基因的序列及其基本信息(表1),主要包括基因信息、亚家族分类、基因位点、系统学命名、别名、蛋白大小和蛋白结构域信息。我们可以获知ABC转运蛋白含有1~2个接合ATP的盒(ABCs)或接合核苷酸的域(nucleotide-binding domains, NBDs)和TMD,具有较高保守性的NBD位于膜的胞质面,能结合并水解ATP供能;而序列相似性较低的TMD能够利用NBF释放的能量选择底物并跨膜转运底物(Theodoulou 2000)。

根据蛋白的大小(全分子、半分子或四分之一分子)、NBDs的结构特征、结构与组织形式和其他生物拟南芥ABC转运蛋白的相似性比较而

表1 拟南芥中ABC转运蛋白基因家族分类以及基因信息

Table 1 The classification and information of ABC transporters families in *Arabidopsis*

基因家族		基因位点	系统学命名	别名	长度/aa	蛋白结构域	蛋白定位*		
亚家族A (subfamily A)	AOH; ABCA	At2g41700.1/2	AtABCA1_1/2		1 882/1 846	(TMD-NBD) <sup>2</sup>	M, C		
		At3g47730.1	AtABCA2	ATATH1	983	TMD-NBD	M, P		
	ATH; ABCA	At3g47740.1	AtABCA3	ATATH2		947		O, P	
		At3g47750.1	AtABCA4	ATATH3		944		M, P	
		At3g47760.1	AtABCA5	ATATH4		872		O, P	
		At3g47770.1	AtABCA6	ATATH5		900		O, P	
		At3g47780.1	AtABCA7	ATATH6		935		O, P	
		At3g47790.1	AtABCA8	ATATH7		901		O, P	
		At5g61730.1	AtABCA9	ATATH11		940		M, P	
		At5g61740.1	AtABCA10	ATATH14		848		O, C	
		At5g61690.1	AtABCA11	ATATH15		954		M, P	
		At5g61700.1	AtABCA12	ATATH16		888		M, P	
		亚家族B (subfamily B)	MDR; DPL (PGP)	At2g36910.1	AtABCB1	PGP1	1 286	(TMD-NBD) <sup>2</sup>	O, P
				At4g25960.1	AtABCB2	PGP2	1 273		S, P
At4g01820.1	AtABCB3		PGP3	1 229		O, V			
At2g47000.1	AtABCB4		PGP4	1 286		O, P			
At4g01830.1	AtABCB5		PGP5	1 230		O, P			
At2g39480.1	AtABCB6		PGP6	1 407		O, Y			
At5g46540.1	AtABCB7		PGP7	1 248		O, P			
pseudogene <sup>e</sup>	AtABCB8				1 241				
At4g18050.1	AtABCB9		PGP9	1 281		O, P			
At1g10680.1	AtABCB10		PGP10	1 227		O, P			
At1g02520.1	AtABCB11		PGP11	1 278		O, P			
At1g02530.1	AtABCB12		PGP12	1 273		O, P			
At1g27940.1	AtABCB13		PGP13	1 245		O, P			
At1g28010.1	AtABCB14		PGP14	1 247		O, P			
At3g28345.1	AtABCB15			1 240		O, P			
At3g28360.1	AtABCB16		PGP16	1 158		O, P			
At3g28380.1	AtABCB17		PGP17	1 240		O, P			
At3g28390.1	AtABCB18		PGP18	1 225		O, P			
At3g28860.1	AtABCB19		PGP19	1 252		O, P			
At3g55320.1	AtABCB20		PGP20	1 408		O, Y			
At3g62150.1	AtABCB21		PGP21	1 292		O, P			
At3g28415.1	AtABCB22			1 221		S, P			
亚家族C (subfamily C)	ATM; DPL (HMT)		At4g28630.1	AtABCB23	ATM1	678	TMD-NBD	M, C	
			At4g28620.1	AtABCB24	ATM2	680		M, C	
	At5g58270.1	AtABCB25	STA1	728		M, M			
	TAP; DPL (TAP)	At1g70610.1	AtABCB26	ATTAP1	700	TMD-NBD	C, C		
		At5g39040.1	AtABCB27	ATTAP2	644		O, E		
	DPL (LLP) MRP; OAD (MRP)	At4g25450.1/2/3	AtABCB28_1/2/3	ATNAP8	714/618/545		M, C		
		At5g03910.1	AtABCB29	ATATH12	634	TMD-NBD	C, C		
		At1g30400.1/2	AtABCC1_1/2	EST1	1 622/1 622	(TMD-NBD) <sup>2</sup>	O, P		
		At2g34660.1	AtABCC2	MRP2	1 623		O, Y		
		At3g13080.1/2/3	AtABCC3_1/2/3	MRP3	1 514/1 489/1 120		O, P		
At2g47800.1		AtABCC4	EST3	1 516		O, P			
At1g04120.1	AtABCC5	MRP5	1 514		S, P				
At3g13090.1	AtABCC6	ATMRP8	1 466		S, P				
At3g13100.1	AtABCC7	ATMRP7	1 493		S, P				
At3g21250.1	AtABCC8	ATMRP6	1 294		C, P				



续表

基因家族	基因位点	系统学命名	别名	长度/aa	蛋白结构域	蛋白定位*	
	At3g60160.1	AtABCC9	ATMRP9	1 490		O, P	
	At3g59140.1	AtABCC10	ATMRP14	1 453		O, C	
	At1g30420.1	AtABCC11	ATMRP12	1 495		O, P	
	At1g30410.1	AtABCC12	ATMRP13	1 495		O, P	
	At2g07680.1	AtABCC13	ATMRP11	1 194		O, P	
	At3g62700.1	AtABCC14	ATMRP10	1 539		O, P	
	At3g60970.1	AtABCC15	ATMRP15	1 037		S, P	
亚家族D (subfamily D)	PMP; FAE	At4g39850.1	AtABCD1	PXA1	1 337	(TMD-NBD) <sup>2</sup>	M, P
		At1g54350.1	AtABCD2		706	TMD-NBD	C, C
亚家族E (subfamily E)	RLI	At3g13640.1	AtABCE1	ATRLI1	603	NBD-NBD	O, Y
		At4g19210.1	AtABCE2	ATRLI2	605		O, Y
		At4g30300.1	AtABCE3	ATNAP15	181		O, N
亚家族F (subfamily F)	GCN; ART (REG)	At5g60790.1	AtABCF1	ATGCN1	595	NBD-NBD	C, N
		At5g09930.1	AtABCF2	ATGCN2	678		C, C
		At1g64550.1	AtABCF3	ATGCN3	715		O, Y
		At3g54540.1	AtABCF4	ATGCN4	723		O, N
		At5g64840.1	AtABCF5	ATGCN5	692		C, C
亚家族G (subfamily G)	WBC; EPD (WHITE)	At2g39350.1	ABCG1		740	NBD-TMD	O, P
		At2g37360.1	ABCG2		755		
		At2g28070.1	ABCG3		730		C, P
		At4g25750.1	ABCG4		577		O, C
		At2g13610.1	ABCG5		649		O, P
		At5g13580.1	ABCG6		727		O, P
		At2g01320.1/2/3/4	ABCG7_1/2/3/4		725/727/728/725		O, P
		At5g52860.1	ABCG8		589		C, P
		At4g27420.1	ABCG9		639		O, P
		At1g53270.1	ABCG10		590		O, P
		At1g17840.1	ABCG11	WBC11	703		O, P
		At1g51500.1	ABCG12	WBC12	687		O, P
		At1g51460.1	ABCG13		678		O, P
		At1g31770.1	ABCG14		648		O, C
		At3g21090.1	ABCG15		691		O, P
		At3g55090.1	ABCG16		720		O, P
		At3g55100.1	ABCG17		662		O, P
		At3g55110.1	ABCG18		708		O, P
		At3g55130.1	ABCG19	ATWBC19	725		O, P
		At3g53510.1	ABCG20		739		O, P
		At3g25620.1	ABCG21		467		C, V
		At5g06530.1/2/3	ABCG22_1/2/3		751/751/691		O, P
		At5g19410.1	ABCG23		624		C, C
		At1g53390.1	ABCG24		1 109		O, C
		At1g71960.1	ABCG25		662		O, P
		At3g13220.1	ABCG26		685		O, P
		At3g52310.1	ABCG27		737		C, P
		At5g60740.1	ABCG28		1 061		S, P
	PDR; EPD (PDR)	At3g16340.1	ABCG29	PDR1	1 416	NBD-TMD-	C, P
		At4g15230.1	ABCG30	PDR2	1 326	NBD-TMD	O, P
		At2g29940.1	ABCG31	PDR3	1 426		O, P
		At2g26910.1	ABCG32	PDR4	1 420		O, P
		At2g37280.1	ABCG33	PDR5	1 413		O, P
		At2g36380.1	ABCG34	PDR6	1 453		M, P

续表

基因家族	基因位点	系统学命名	别名	长度/aa	蛋白结构域	蛋白定位*	
	At1g15210.1	ABCG35	PDR7	1 442		C, P	
	At1g59870.1	ABCG36	PEN3	1 469		C, P	
	At3g53480.1	ABCG37	PDR9	1 450		O, P	
	At3g30842.1	ABCG38	PDR10	1 406		O, P	
	At1g66950.1	ABCG39	PDR11	1 454		C, P	
	At1g15520.1	ABCG40	PDR12	1 423		C, P	
	At4g15215.1	ABCG41	PDR13	1 390		O, P	
	At4g15233.1	ABCG42		1 170		O, P	
	At4g15236.1	ABCG43		1 388		O, P	
亚家族I (subfamily I)	CCM	At1g63270.1	AtABC11	ATNAP10	229	NBD	M, C
		Atmg00110.1	AtABC12		256	TMD	
		Atmg00900.1	AtABC13		256		
		At2g07681.1	AtABC14		256		M, P
		At2g07771.2	AtABC15_2		256		M, P
	ISB	At3g10670.1	AtABC16	ATNAP7	338	NBD	C, C
		At1g32500.1	AtABC17	ATNAP6	475	CYT	C, C
		At4g04770.1	AtABC18	LAF6	557		C, C
		At5g44316.1	AtABC19		470		C, C
	CBY (Y179)	At4g33460.1	AtABC110	EMB2751	271	NBD	C, C
		At5g14100.1	AtABC111	ATNAP14	176		C, C
	MKL; TGD	At3g21580.1/2	AtABC112_1/2		205/384	TMD	C, P
		At1g65410.1	AtABC113	TGD3	345	NBD	C, C
		At1g19800.1/2/3	AtABC114_1/2/3	TGD1	350/350/350	TMD	O, P
	NO	At3g20320.1/2	AtABC115_1/2	TGD2	381/282	SSA	C, C
		At2g37300.1	AtABC116		128	TMD	O, N
		At1g67940.1	AtABC117	ATNAP3	263	NBD	O, G
	NO (ADT)	At1g03900.1	AtABC118	ATNAP4	272		O, N
		At1g03905.1	AtABC119		290	NBD	O, N
		At5g02270.1	AtABC120	ATNAP9	328		O, Y
At5g44110.1/2/3		AtABC121_1/2/3	POP1	282/214/223		O, N	
其他(others)	At1g71330.1	ATNAP5				O, P	
	At2g37010.1	ATNAP12				S, P	

\*代表预测的蛋白双定位情况, C代表叶绿体, M代表线粒体, S代表分泌蛋白, P代表膜蛋白, O代表未知。NBD: 核苷酸结合结构域(ATP结合结构域); TMD: 跨膜结构域; SSA: 基质结合蛋白; CYT: 与ABC结构域相互作用的保守的可溶性蛋白; CCM: 起源于细胞色素C的家族; ISB: 起源于铁硫中心的家族; CBY: 类似于假设的钴摄取系统的家族; Y179: 类似于*Methanococcus janaschii* Y179蛋白的亚家族; MKL: 类似于*Mycobacterium leprea* MKL蛋白的家族; TGO: 三半乳糖二酰基甘油; NO: 未知功能的蛋白, 表面上与已知的家族不相关; ADT: 未知功能的蛋白, 来自一个同构集群。

分为8个亚家族拟南芥(亚家族A~G、I)和其他两个成员(ATNAP5和ATNAP12)。其中最大的亚家族为亚家族G, 成员多达43个, 其次为亚家族B和I。在各个亚家族之间ABC转运蛋白的结构域组织形式多种多样; 通常, 全分子的ABC转运蛋白包含2个NBD和2个TMD, 4个结构域结合在一起才能执行转运功能。半分子的ABC转运蛋白通过形成二聚体或多聚体来实现功能, 这种二聚体可为同二聚体, 又可为异二聚体。通常情况下, NBD和TMD存在于不同的多肽上的转运蛋白称为1/4分子转运

蛋白。亚家族G有15个全分子蛋白和28个半分子蛋白; 亚家族B有22个全分子蛋白和7个半分子蛋白; 亚家族I全为四分之一分子蛋白。拟南芥基因组编码的131个ABC转运蛋白的长度分布范围广泛, 蛋白分子大小从250到1 800个氨基酸不等。最短的蛋白质为四分之一分子蛋白AtABC116, 仅含128个氨基酸, 而最长的则为全分子蛋白AtABC11\_1, 含有多达1 882个氨基酸。

此外, 我们还通过蛋白定位预测发现极少数的蛋白为分泌蛋白, 大部分蛋白质定位于叶绿

体、线粒体和膜蛋白上,说明这些蛋白质与物质的跨膜运输有关。为了进一步了解ABC转运蛋白的生物学功能,我们还通过GO功能分类进行富集分析。结果(图1)表明,在生物学过程中,在获取的

128个基因GO中,ABC转运蛋白比较集中于各类物质的运输功能(transport),除此外,还包括对药物反应(response to drug)和生物刺激反应(response to biotic stimulus)等等。总之,各ABC转运蛋白之间

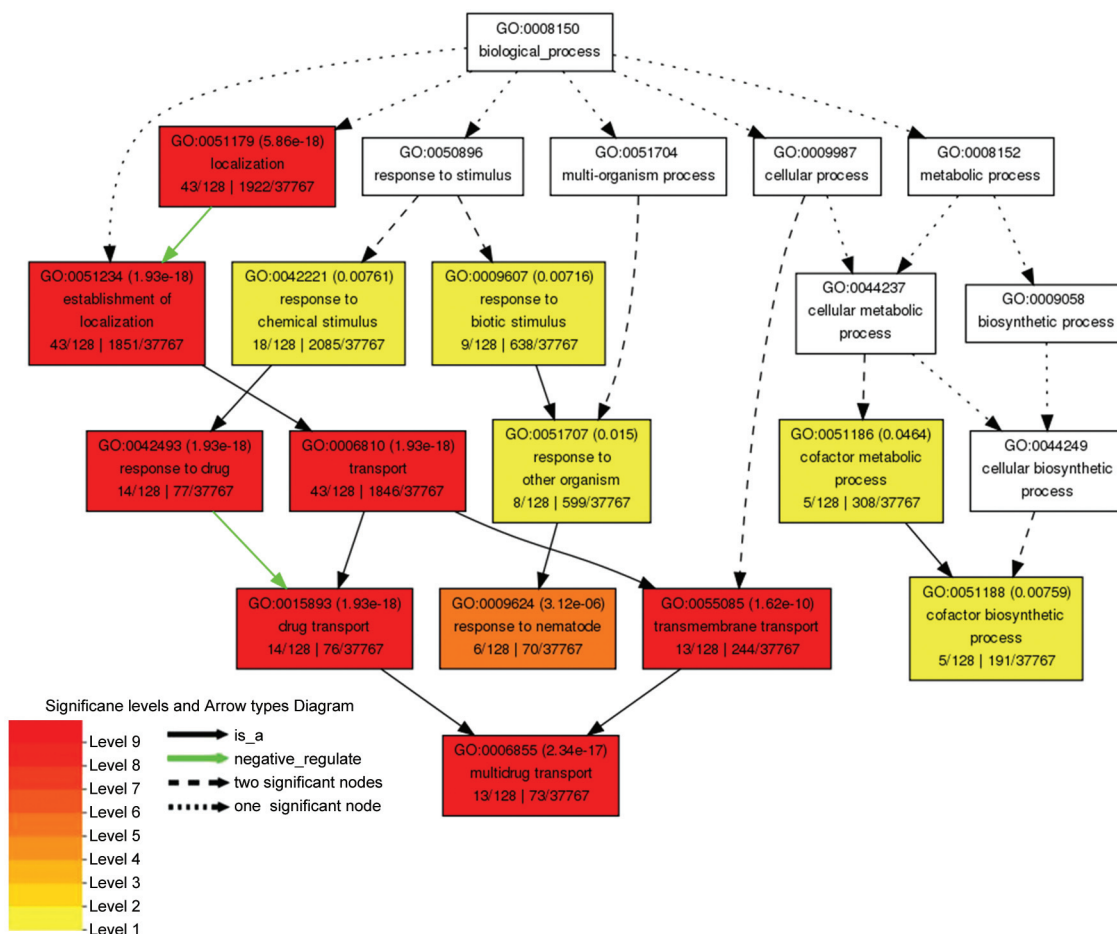


图1 拟南芥ABC转运蛋白GO生物学过程富集分析

Fig.1 *Arabidopsis thaliana* ABC transporter protein GO (Gene Ontology) biological process enrichment analysis

的特定区域的范围内氨基酸序列上虽有同源性,但总体结构差异巨大,这种差异可能与功能的差异有关。

## 2 拟南芥ABC转运类蛋白在染色体分布及复制事件

为了确定ABC转运蛋白基因在染色体的分布情况,我们使用染色体图工具,将预测的所有ABC转运蛋白基因定位到5条染色体上的相应位置,并分析相应的基因家族连锁信息。结果(图2)表明,129个基因家族成员(除两个位于线粒体基因上的家族成员外)在拟南芥第1至5染色体上都有分布,其中第1和3条染色体数目最多,分别有31和40个

基因家族成员,而第2、4和5染色体分布相比较少,分别为20、18和20个基因(图2)。所有的数染色体两臂上均有分布,但从图2可以看出,拟南芥的ABC转运蛋白基因在各个染色体上并非均匀分布,某些区域的分布密度很大,比如,拟南芥3号染色体有44个基因分布,这44个基因均集中在5个较小的区域范围内,尤其下半部聚集较小的2个区域内就聚集了25个基因,而物理长度长于3号染色体的5号染色体上半部总共也只分布7个基因。而且结果还表明很多基因是以基因簇的方式存在,尤其是存在2个基因紧密连锁关系,从图2中可以看出,

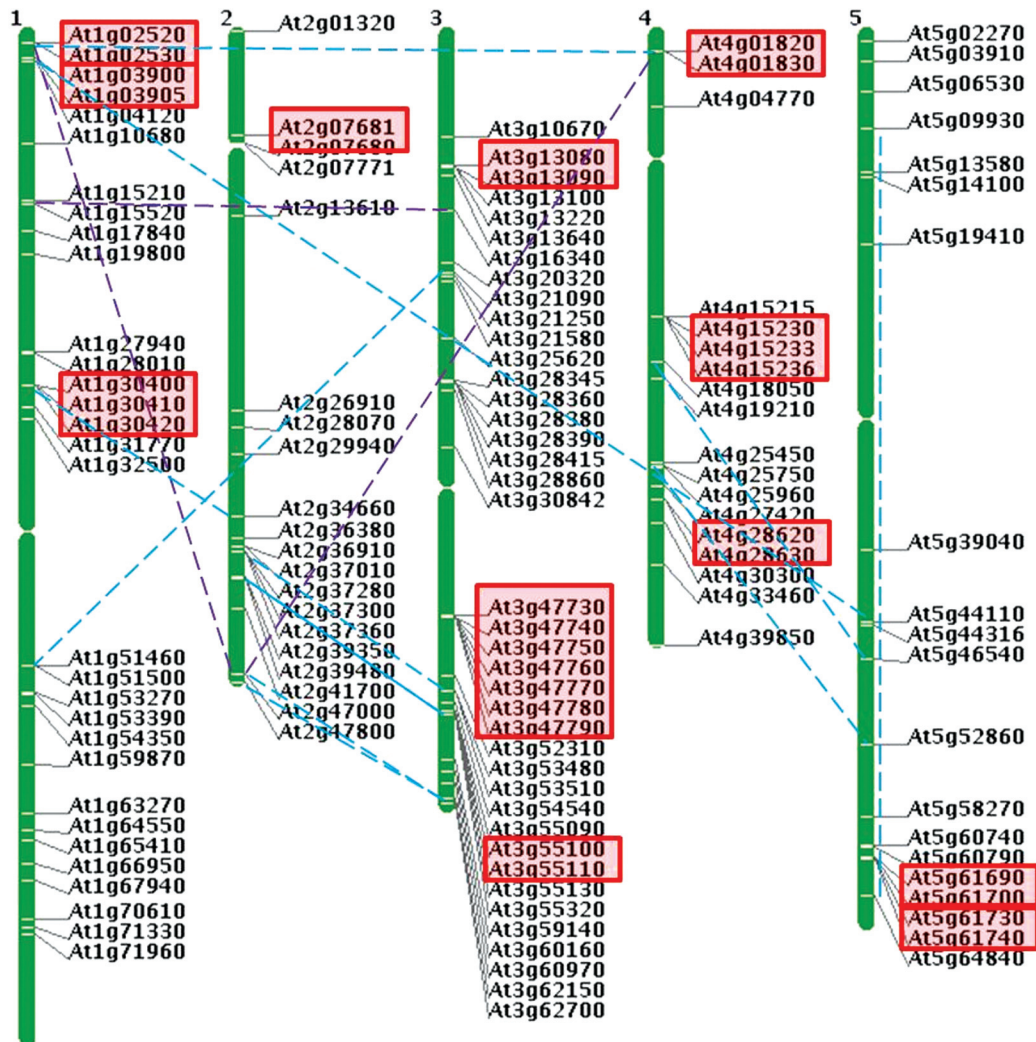


图2 拟南芥ABC转运蛋白基因的染色体分布和基因复制示意图

Fig.2 The sketch maps of *Arabidopsis* ABC genes on different chromosomes and gene duplication  
基因位点用短横表示, 基因簇聚区(连续基因)用方框表示, 虚线代表基因之间在进化过程中的可能复制关系。

至少有12个基因簇存在, 每个基因簇包括至少2个基因是紧密连锁的, 尤其3号染色体的下半部分就有7个ABCA家族的基因(At3g47730、At3g47740、At3g47750、At3g47760、At3g47770、At3g47780和At3g47790)紧密连锁, 说明他们在进化过程中有可能起源同一片段复制和串联重复导致。

目前的研究表明, 拟南芥的进化过程中, 至少发生了1次或3次多倍化事件, 大多数基因都有两个甚至更多拷贝(Ermolaeva等2003)。因此, 确定染色体同源片段是基因组学研究的一个重要方面, 有助于揭示基因组在历史上发生的多种多样的进化事件。为进一步研究基因复制与拟南芥ABC转

运类基因进化的关系, 我们在了解染色体分布的基础上, 又利用TAIR上已有的染色体片段信息, 对比分析公共资源数据库Paralogous in *Arabidopsis thaliana*基因组区域复制信息, 来寻找ABC转运类基因可能的复制进化事件或基因来源信息。结果(图2)表明, 在拟南芥ABC转运类131个基因中, 在已知的DNA大片段复制区域找到15对ABC转运类旁系同源基因(在图2中用虚线连接), 这些基因具有序列同源性高特点, 且都分别处于已经发现的基因簇聚区。其中12对ABC转运基因来源是发生在较近年代(recent age)的复制事件过程中, 分别是At1g02520与At4g01820, At1g03900与



At5g44110, At1g30400与At2g34660, At1g51500与At3g21090, At2g37280与At3g53480, At2g39350与At3g55130, At2g39480与At3g55320, At2g47000与At3g62150, At2g47800与At3g62700, At4g18050与At5g46540, At4g25750与At5g52860, At5g09930与At5g64840, 还有3对ABC转运基因来源是发生在较早年代(old age)复制事件过程中, 分别是At1g02520与At2g47000, At1g15210与At3g16340, At2g47000与At4g01830。因此, 通过对拟南芥ABC转运类基因功能和进化关系复制的分析, 不但说明拟南芥中ABC转运蛋白重复基因中的一部分基因是来源于复制事件, 而且对植物染色体进化和基因组的多倍化事件也有重要的借鉴作用。

### 3 拟南芥ABC转运类蛋白组织器官表达模式分析

同源基因或结构序列相近基因在不同组织器官中的表达模式分析, 能为这些基因的生理生化功能研究提供了重要信息。因此, 为了阐明拟南芥ABC转运蛋白在植物生长发育过程中的功能, 本文利用生物信息学方法, 通过对拟南芥数据库的芯片数据进行收集, 聚类分析拟南芥ABC转运蛋白在各个组织器官和不同发育时期中的mRNA丰度。除了其中9个基因(AtABCA8/At3g47790.1、AtABCB8/pseudogene、AtABCB22/At3g28415.1、ABCG38/At3g30842.1、ABCG41/At4g15215.1、ABCG42/At4g15233.1、ABCG43/At4g15236.1、AtABCII1/At1g63270.1和AtABCII9/At5g44316.1)因缺少芯片探针的原因, 没有收集到相应的表达谱芯片数据外, 我们获取了其他122个基因在47个组织器官和不同发育时期的基因表达数据。研究结果(图3)首先表明, 拟南芥122个ABC转运蛋白mRNA在各个组织和时期中都有表达, 进一步证明ABC转运蛋白是一类比较重要功能蛋白, 其可能参与到植物生理发育的各个过程。

其次, 为了更加细致和全面的了解ABC转运蛋白在不同组织部位以及发育时期的具体功能, 以及了解各个亚家族之间功能的异同, 我们又根据这些基因在各个组织和发育过程中表达的差异进行层次聚类分析, 结果(图3)可以把这8个亚家族中的122个基因可以分成9个小组(I~IX), 每个小组都是3至4个亚家族组成。在不同组织器官和发育时期各个基因亚家族存在较大差异。

I组共计10个基因, 主要包含ABCG (21、29、23和25这4个成员)、ABCB (1、16、20和6这4个成员)以及ABCA (1.1和1.2这2个成员)三类家族, 主要表达特点是表达模式共性较为集中在茎节的第二节点(图3, 37)高表达, 但他们之间也存在较大差异, 比如, ABCG的4个成员还在茎节的第一节点(图3, 3)特异性的较高表达, 而ABCB和ABCA这两个家族在吸胀的种子(图3, 2)和衰老的叶片(图3, 33)中有特异性的较高表达, 说明这三个基因亚家族存在共同参与的生理过程, 如茎节第二节点的发育, 但也存在各自参与的不同组织和时期的发育。

II组共计15个基因, 主要包含ABCG (9、14、34、37、35、15、30和3这8个成员)、ABCB (11、12、17、4和15这5个成员)、ABCC6和ABCF3四类家族, 主要表达特点是表达模式共性较为集中在根(图3, 7和26)和吸胀的种子(图3, 2)中。此外ABCB11和12还在拟南芥花器官15期的心皮(图3, 16)和种子3期/果荚(图3, 39)中有较高水平的表达, 说明这两个基因除了参与根的发育以外, 可能还参与了种子形成的过程。

III组共计20个基因, 主要包含ABCI (10、14.1、14.2、14.3、11、15.1、15.2、12.1、12.2、6和7这11个成员)、ABCB (21、23、24、14和29这5个成员)、ABCG (36和24这2个成员)、ABCC13和ABCD2这五类家族, 主要表达特点是表达模式共性较为集中在幼嫩的叶片(图3, 21)和莲座叶(图3, 47)中。不过整体来看该亚组的基因其表达范围非常广泛, 但表达量相对不高, 其表达范围囊括了种子、叶片、根、花和果荚, 但是在雄蕊中表达的基因很少, 说明此类基因可能主要参与了拟南芥营养生长的整个过程, 对生殖生长过程的参与较少。

IV组共计23个基因, 主要包含ABCC (1.1、1.2、11、12、8、10、15、9、7、4和2这11个成员)、ABCG (1、19、40、22.1、22.2、22.3和27这7个成员)、ABCB (27和26这2个成员)、ABCI (17和8这2个成员)和ABCF5这五类家族, 主要表达特点是表达模式共性较为集中在花器官15期的萼片(图3, 18)和枯萎的叶片(图3, 33)中。此外ABCG1、19、40和ABCC1.1、1.2在花器官15期

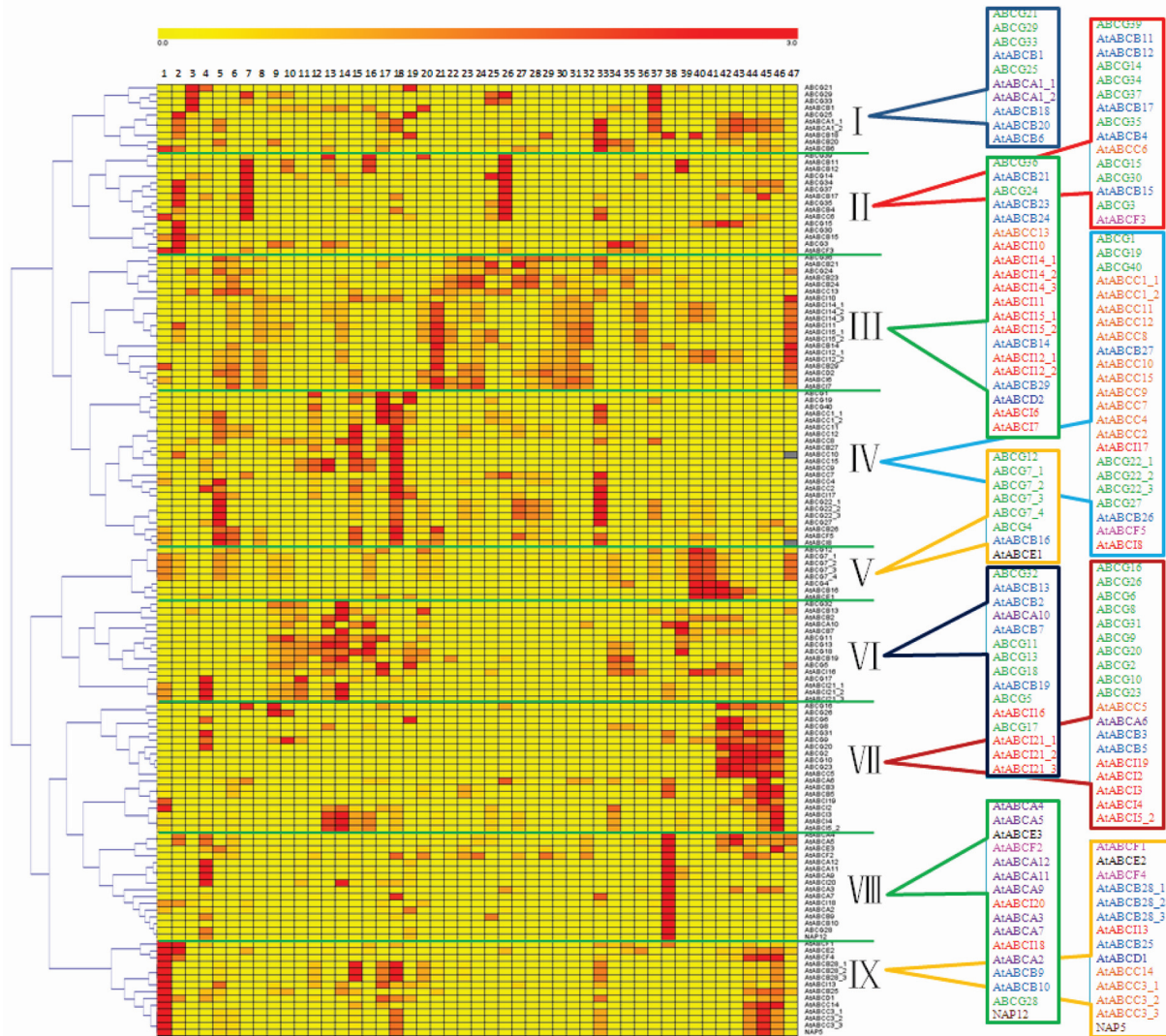


图3 拟南芥ABC转运蛋白基因在组织器官表达模式的系统聚类

Fig.3 Hierarchical cluster display of expression profiles for *Arabidopsis* ABC genes in different tissues and/or organs

颜色标尺代表表达值以2为底的对数。1: 干种子; 2: 吸胀的种子(24 h); 3: 茎的第一节点; 4: 第12期花的雄蕊; 5: 茎叶; 6: 子叶; 7: 根; 8: 过渡期后到开花的整个莲座丛; 9: 第9期花; 10: 第10和11期花; 11: 第12期花; 12: 第15期花; 13: 第12期花的心皮; 14: 第12期花的花瓣; 15: 第12期花的萼片; 16: 第15期花的心皮; 17: 第15期花的花瓣; 18: 第15期花的萼片; 19: 第15期花的雄蕊; 20: 第15期花的花梗; 21: 叶1+2; 22: 7叶叶柄; 23: 7叶叶远侧半; 24: 7叶叶近端半; 25: 下胚轴; 26: 根; 27: 莲座叶2; 28: 莲座叶4; 29: 莲座叶6; 30: 莲座叶8; 31: 莲座叶10; 32: 莲座叶12; 33: 衰老的叶片; 34: 茎尖花序; 35: 茎尖过渡; 36: 茎尖营养; 37: 茎节(第2节间); 38: 成熟的花粉粒; 39: 种子阶段3 w/果荚; 40: 种子阶段4 w/果荚; 41: 种子阶段5 w/果荚; 42: 种子阶段6 w/果荚; 43: 种子阶段7 w/果荚; 44: 种子阶段8 w/果荚; 45: 种子阶段9 w/果荚; 46: 种子阶段10 w/果荚; 47: 莲座叶。

的花瓣(图3, 17)中表达量较高, 大部分基因在花器官12期的萼片(图3, 15)和茎叶(图3, 5)中也有所表达。说明此类基因主要参与花瓣与叶片的发育。

V组共计8个基因, 主要包含ABCG (12、7.1、7.2、7.3、7.4和4这6个成员)、ABCB16和ABCE1这三类家族, 主要表达特点是表达模式共性较为集中在种子4期/果荚(图3, 40)和种子5期/果

荚(图3, 41)中, 随着种子的发育, 参与的基因逐渐减少。此类基因的表达较为特异, 说明此类基因主要参与种子(果荚)的发育。

VI组共计15个基因, 主要包括ABCG (32、11、13、18、5和17这6个成员)、ABCB (13、2、7和19这4个成员)、ABCI (16、21.1、21.2和21.3这4个成员)和ABCA10这四类家族, 主要表达特点



是表达模式共性较为集中在花器官12期的花瓣(图3, 14)中, 此外一些基因也在心皮中有所表达, 说明此类基因主要参与花器官发育的过程。

VII组共计19个基因, 主要包括ABCG (16、26、6、8、31、9、20、2、10和23这10个成员)、ABCI (19、2、3、4和5.2这5个成员)、ABCB (3和5这2个成员)、ABCC5和ABCA6这五类家族, 主要表达特点是表达模式共性较为集中在种子6~10期/果荚(图3, 42~46)中, 说明此类基因参与了拟南芥种子发育的6到10期, 对于种子的形成有重要的作用。

VIII组共计16个基因, 主要包括ABCA (4、5、12、11、9、3、7和2这8个成员)、ABCI (20和18这2个成员)、ABCB (9和10这2个成员)、ABCE3、ABCF2、ABCG28和NAP12这七类家族, 主要表达特点是表达模式共性主要集中在成熟的花粉粒(图3, 38)中, 部分基因(ABCA11、12、9和ABCI20)在花器官12期的花药(图3, 4)中表达量较高。这说明此类基因主要参与拟南芥花粉粒的发育过程。

IX组共计14个基因, 主要包括ABCB (28.1、28.2、28.3和25这4个成员)、ABCC (14、3.1、3.2和3.3这4个成员)、ABCF (1和4这2个成员)、ABCE2、ABCD1、ABCI13和NAP5这七类家族, 主要表达特点是表达模式共性主要集中在干燥的种子(图3, 1)中, 此外在种子8-10期/果荚(图3, 44~46)中也有所表达。这说明该类基因主要参与了种子的发育, 尤其是成熟种子的形成。

综合以上基因表达模式和聚类分析说明: 拟南芥ABC基因家族基因的表达在进化过程中已经发生了显著的功能分化, 各个亚家族之间的表达模式的存在较大的差异, 存在的功能多样性, 但并不是完全随机分布, 仍具有一定的规律性, 说明彼此之间功能的具有一定的协调性和严格调控。

#### 4 拟南芥ABC转运类蛋白在花药发育中的功能分析

为进一步说明ABC转运蛋白在正常生理过程中存在的功能多样性和协调性, 我们选择了花药发育过程这一代谢特殊过程来分析ABC转运蛋白的功能。花药形成是一个十分复杂的过程, 其中花粉囊壁的最内一层细胞绒毡层的功能和作用至关重要。绒毡层细胞具有高度的代谢活性, 含有

较多的RNA、蛋白质和酶、丰富的油脂、类胡萝卜素和孢粉素等物质。这些物质对花粉粒的发育形成起着重要的营养和调节作用, 为花粉壁的建设提供前体物质。而绒毡层细胞作为一种分泌型细胞, 其合成的物质需要从细胞中运输出来, 才能分泌到花药腔内从而发挥作用(Zhang等2011)。绒毡层细胞的质膜上分布着各种跨膜蛋白通道, 保证相应的底物能够运输出来。目前已经有研究表明, ABC转运蛋白参与这个运输过程, 比如ABCG26/WBC27可能参与孢粉素前体的运输, *abcg26*突变体无法形成可育的花粉粒, 花粉外壁在后期降解(Choi等2011; Quilichini等2010; Xu等2010)。

我们通过提取花药发育的各个时期的表达谱芯片和聚类分析, 来了解和分析ABC转运蛋白在花药发育中的功能; 主要选取绒毡层代谢物质形成、发育和功能最主要的时期, 对应到花的发育时期分明别是: S9 (早期)、S10/11和S12 (中期)、S15 (后期)以及特异性器官S12-S (中期的雄蕊)和MP (成熟的花粉粒)。研究结果(图4)表明, 根据ABC转运蛋白基因在不同时期和组织之间的特异性表达情况, 主要可以分为6大类:

第一类(图4-A)中包含11个基因, 主要包括ABCG (9、39)、ABCA (10)、ABCB (7、2611、12和15)、ABCI (12.1和12.2)和ABCC6, 部分此类基因主要在成熟的花粉粒中和S10/11期表达, 表达特异性不强, 且表达量并不高, 说明该类基因可能是协助作用, 主要可能参与了花粉粒的中期和后期的成熟过程。第二类(图4-B)中包含50个基因, 此类基因主要在S9和S10/11期表达, 部分在S15期表达, 也具有表达特异性不强和表达量也不高的特点, 但在成熟花粉粒中没有任何表达(与第一类的主要区别), 说明此类基因也是协助作用, 但主要的功能是花发育的早前。第三类(图4-C)中包含33个基因, 主要包括ABCA、ABCB和ABCG类, 它们的共同特点是都在雄蕊和成熟花粉粒中表达, 且大部分基因表达量较高, 约一半的基因在雄蕊中也有所表达, 说明此类基因参与了拟南芥花药的发育与成熟花粉粒的形成。第四类(图4-D)中仅包含8个基因, 但此类基因的表达主要集中在S9期, 且具有较高表达量以及较特异的特性, 而在在雄

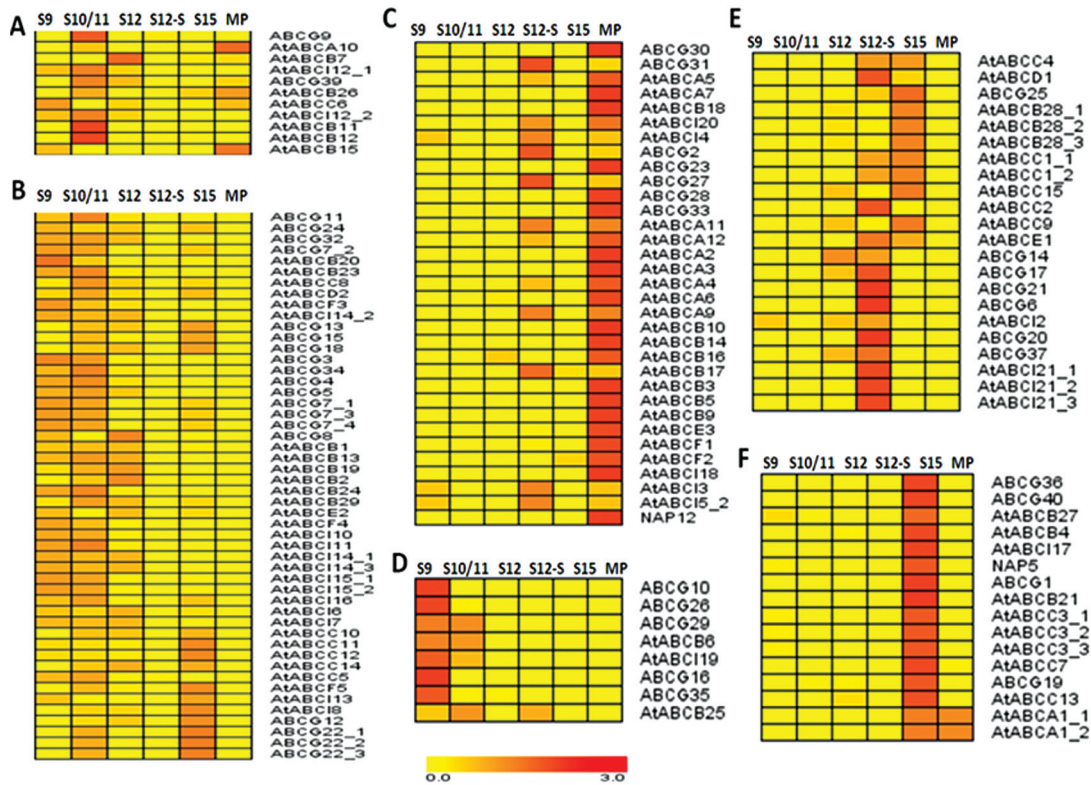


图4 拟南芥ABC基因在花药发育过程中的系统聚类

Fig.4 Hierarchical cluster display of expression profiles for *Arabidopsis* ABC genes during the anther development  
颜色标尺代表表达值以2为底的对数。

蕊和成熟花粉粒中基本没有表达,说明此类基因可能不参与拟南芥花药的发育。第五类(图4-E)中包含22个基因,主要包括ABCB、ABCC和ABCG类,此类基因大部分主要在花药中表达,但是在成熟的花粉粒中没有表达(与第三类的主要区别),说明此类基因参与了拟南芥花药的发育。第六类(图4-F)中包含16个基因,主要包括ABCC和ABCG类基因,非常特异和较高表达在S15期,此类基因在花药中没有表达,只有两个基因在成熟的花粉粒中表达量较高,说明此类基因可能不参与拟南芥花药的发育。综合以上所述,大部分ABC转运蛋白在花药不同的发育时期和组织部位都有表达变化,说明功能的广谱性,但又有各自的时间差异和组织表达的特异性,说明彼此功能分工和协调具有严格的调控。

### 5 拟南芥ABC半分子转运类蛋白相互作用网络分析(蛋白互作图谱与关键基因)

在前面的基因信息分析中,我们知道在ABC

蛋白家族里属于ABC半分子转运蛋白的有5个亚家族: PMP (2个成员)、WBC (29个成员)、ATH (16个成员)、ATM (3个成员)和TAP (2个成员);作为“半转运蛋白”的ABC蛋白主要是按照反方向排列的(NBF)-(TMD)型转运蛋白,它们必须形成同源或异源二聚体(NBF-TMD)-(NBF-TMD)结构来发挥功能。每个TMD结构由6个跨膜 $\alpha$ 螺旋(TM<sub>H</sub>s)组成,两个TMD形成一个运输物质的通道结构。因此,为了进一步了解和分析这些ABC半分子转运蛋白可能存在的相互关系,我们进一步对ABC转运蛋白之间进行蛋白相互作用(PPI)分析,结果(图5)表明,对于整个ABC转运蛋白家族来说,共获得719节点和1 015个边线,其中,主要节点中心基因大部分属于ABCG、ABCA、ABCB和ABCC这4个亚家族,分别有14、3、16和7个基因节点。从蛋白互作图谱上可以看出,通过预测发现ABCG11与ABCG12有相互作用,而已经发表的文献也证实了ABCG11确实可以与ABCG12形成异源二聚体



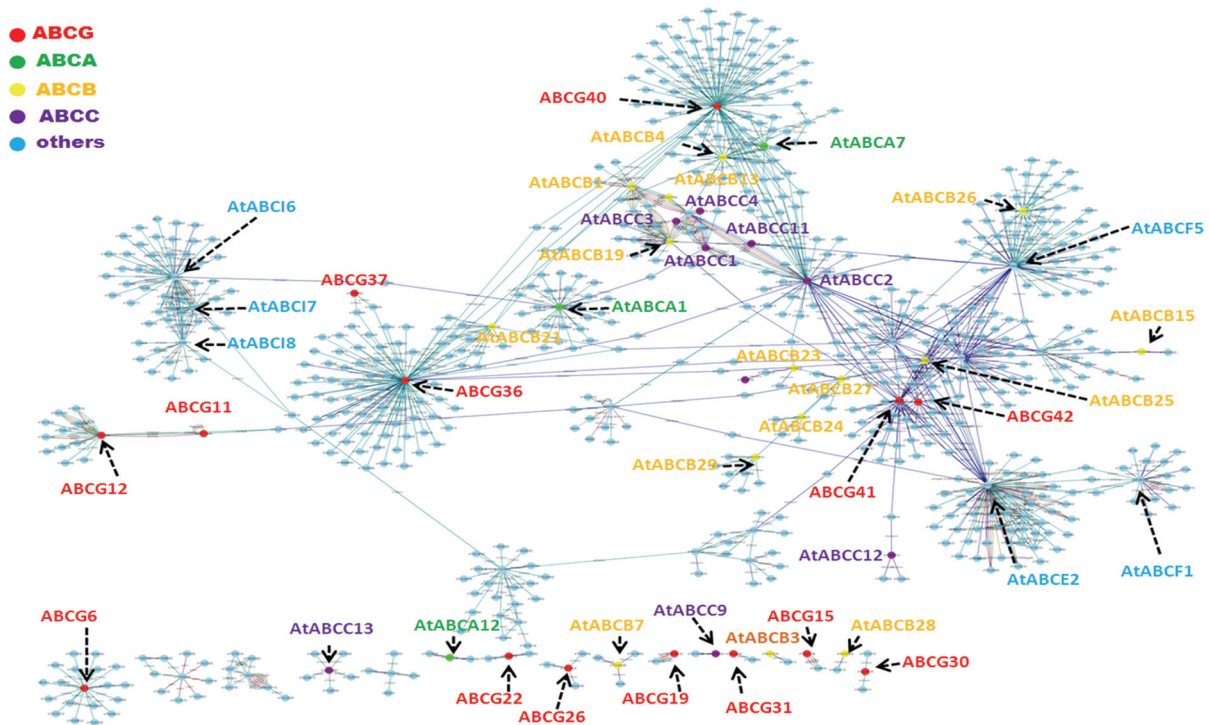


图5 拟南芥ABC半分子转运类蛋白的相互作用预测分析

Fig.5 The *Arabidopsis* ABC half-transporter protein-protein interaction (PPI) analysis

从而发挥功能(McFarlane等2010),这说明我们的预测具有一定的可信度,可为今后的蛋白互作研究提供方向。我们可以看出,异源二聚体的形成,不但存在于各个亚族内部之间(作用位点较少,图1),而且更多的是存在于各个亚族之间(图5);比如,ABCG亚家族的ABCG42不但与同属亚家族的ABCG41相互作用,而且还与ABCB25亚家族的存在相互作用,从而增加了功能的多样性以及控制转运底物协助的准确性。进一步结合这些基因的时空分析(表达谱数据),就能全面的分析和预测这些半分子之间存在相互作用的时间和部位,从而更好的解释这些基因的关系和可能的生物功能。

## 讨 论

很多研究已经表明,在进化中由于基因组复制存在,植物基因家族成员的重复数目常常多于相应的动物家族,这些基因众多拷贝使得植物具有适应环境变化的遗传多样性(Xu等2012)。ABC转运蛋白基因也不例外,拟南芥和水稻基因组各编码130个以上的ABC蛋白,两倍于人类、果蝇和线虫(50~70个成员),约占转运蛋白开放阅读框总

数的1/6,反映了该家族在植物生长发育过程中的重要作用。ABC转运蛋白广泛存在于各类植物中,参与了植物生长发育的各个过程,但是对于植物中众多的ABC转运蛋白基因数目,目前已经报道的基因功能研究并不多见。尽管植物的ABC转运蛋白结构的相似性以及NBD的高度保守性,不同亚簇ABC转运蛋白的作用机制具有一定的相似性,但生物学过程中运输底物的运输需要很多转运蛋白的参与,对于每种底物所对应的一种或几种转运蛋白,从而造成ABC转运蛋白的功能多样性、不确定性和复杂性;目前已知植物体内ABC转运蛋白及其底物的研究主要包括:拟南芥ABCG12(CER5)能够从表皮细胞转运超长链脂肪酸形成蜡质前体,如果破坏其活性会造成蜡质的非正常积累。在*abcg12*(*cer5*)突变体中,其茎表面的蜡质含量减少,但是在*abcg12*中表皮的蜡质总量并没有减少,说明突变体只是破坏了蜡质单体的运输,并没有影响到蜡质的合成(Pighin等2004)。ABCG11(DSO)能够运输长链和超长链脂肪酸(如脂质和蜡质的前体物质),*abcg11*(*dso*)表现出表皮蜡质的减少,并且没有正常形态的表皮细胞和气孔。*dso*蜡

质与角质组成发生变化,说明ABCG11运输蜡质和角质单体,对形成表皮尤为重要(Bird等2007; Panikashvili等2007; Ukitsu等2007)。ABCG19可以使烟草具有卡那抗性(Mentewab和Stewart 2005)。最近,人们发现ABCG25形成同源二聚体并且可以介导脱落酸(ABA)的运输,从而使植物可以抵御干旱的环境(Kang等2010; Kuromori等2010)。ABCG26/WBC27可能参与孢粉素前体的运输, *abcg26*突变体无法形成可育的花粉粒,花粉外壁在后期降解,并且绒毡层内壁没有正常乌氏体的累积(Choi等2011; Quilichini等2010)。此外,棉花中的ABCG1 (GhWBC1)和烟草中的ABCG1 (NtWBC1)也被鉴定出来。GhWBC1推测与纤维素的伸长有关,因为*ghwbc1*突变体中的纤维素显著变短, *GhWBC1*在拟南芥中过量表达会造成角果缩短(Zhu等2003)。NtWBC1在柱头特异表达,说明可能与生殖过程相关,但是还需要更深一步的研究(Otsu等2004)。一些植物ABCG转运蛋白的底物和动物的某些ABCG蛋白一样具有亲脂性。在人体内,ABCG1可以运输胆固醇和磷脂质(Klucken等2000), ABCG2则可以运输类固醇(Janvilisri等2003)。而且,人类的ABCG5和ABCG8可以形成异源二聚体从而运输植物甾醇(Berge等2000)。在植物的表面覆盖着脂质从而形成一层保护膜使植物抵抗外界干旱,外孢壁(花粉外壁)是其中的一种。最外层的表皮细胞覆盖了植物花药腔内所有的器官,并且使植物抵抗各种环境的侵害。这层壁的细胞可以合成形成表面角质的组分。角质是不同的C<sub>16-18</sub>脂肪酸组成的聚合物(Kolattukudy 2001)。在大多数开花植物中,角质中中链羟基化脂肪酸占优势,这被认为对于形成延长二维聚合物单层是必要的(Heredia-Guerrero等2009)。因此,这些底物的运输需要很多转运蛋白的参与,对于每种底物所对应的一种或几种转运蛋白将成为未来的研究重点。

研究模式生物拟南芥的ABC转运蛋白对于揭示此类蛋白在植物体内的生理过程的分子机制是十分具有借鉴意义和参考价值的。本研究中,系统的分析和介绍了拟南芥ABC转运蛋白主要有8个亚家族拟南芥(亚家族A-G、I)和其他两个成员(ATNAP5和ATNAP12)组成,获知ABC转运蛋白的

结构域组织形式多种多样,在各个亚家族之间存在着差异。其成员既包括跨膜的转运蛋白也包括不含跨膜结构域的可溶性蛋白,在物质的跨膜运输以及蛋白质翻译的调控中发挥重要作用。此外,ABC转运蛋白之间的氨基酸序列上虽有同源性,但结构差异巨大,这说明这些基因在复制以后,结构域发生变化,对基因产物的功能产生影响,逐渐形成一些特有的功能,这对蛋白质组多样性有重要意义。

基因复制(gene duplication)是基因家族扩张过程中的重要事件,复制的基因为新基因的产生和功能分化提供了物质前提。基因复制发生主要包括:片段复制(segmental duplication)、串联重复(tandem repeat)和逆转录转座(retro transposition)事件这3种机制。在本研究中,ABC转运蛋白在拟南芥基因组找到了至少15对片段复制现象,分别是B家族的At1g02520与At4g01820, At2g39480与At3g55320, At2g47000与At3g62150, At4g18050与At5g46540, At1g02520与At2g47000, At2g47000与At4g01830; C家族的At1g30400与At2g34660, At2g47800与At3g62700; F家族的At5g09930与At5g64840; G家族的At1g51500与At3g21090, At2g37280与At3g53480, At2g39350与At3g55130, At4g25750与At5g52860, At1g15210与At3g16340; I家族的At1g03900与At5g44110。其中以B家族的片段复制事件最多。这说明这些基因在复制以后,结构域发生变化,对基因产物的功能产生影响,逐渐形成一些特有的功能,这对蛋白质组多样性有重要意义。因此,拟南芥ABC转运类基因功能和进化复制研究,对拟南芥染色体进化和基因组的多倍化事件有重要的借鉴作用。

对同一家族基因在不同组织器官中的表达模式分析,能为这类基因的生理生化功能研究提供重要信息。本研究利用生物信息学方法,对拟南芥数据库的芯片数据进行收集和聚类分析,分析131个ABC转运基因在47个组织器官和不同发育时期中的基因表达情况,结果表明,在拟南芥中任何特定的生理过程都至少需要一个ABC转运蛋白(图3),大部分ABC转运蛋白在不同的发育时期和组织部位都有表达变化。为进一步说明ABC转运蛋白在正常生理过程中可能存在的功能的多样

性和协调性, 我们重点分析了代谢过程比较旺盛的花药发育过程来分析ABC转运蛋白的表达变化, 结果同样表明, 拟南芥ABC转运蛋白家族基因的表达在进化过程中已经发生了显著的功能分化, 各个亚家族之间的表达模式的存在较大的差异, 但并不是完全随机分布, 仍具有一定的规律性, 说明拟南芥ABC转运蛋白功能的广谱性, 但又有各自的时间差异和组织表达的特异性, 说明彼此功能分工和协调不同。

蛋白质相互作用网络系统生物学的一个重要方面。了解植物蛋白质相互作用网络和相互作用组将提供理解调控植物发育、生理和病理过程的关键信息。且在ABC蛋白家族里属于ABC半分子转运蛋白必须形成同源或异源二聚体才能相互作用。本研究利用拟南芥ABC转运蛋白互作图谱, 我们系统的整理分类ABC转运蛋白的关系, 从而可以用于揭示协同作用蛋白的系统网络。ABC转运蛋白利用同源或异源二聚体的相互作用, 从而增加了功能的多样性以及控制转运底物协助的准确性。进一步结合基因的表达谱数据和蛋白相互作用, 就能全面的分析和预测这些半分子之间存在相互作用的时间和部位, 从而更好的解释这些基因的关系和可能的生物功能。

### 参考文献

- Berge KE, Tian H, Graf GA, Yu L, Grishin NV, Schultz J, Kwitrovich P, Shan B, Barnes R, Hobbs HH (2000). Accumulation of dietary cholesterol in sitosterolemia caused by mutations in adjacent ABC transporters. *Science*, 290: 1771~1775
- Bird D, Beisson F, Brigham A, Shin J, Greer S, Jetter R, Kunst L, Wu XW, Yephremov A, Samuels L (2007). Characterization of *Arabidopsis* ABCG11/WBC11, an ATP binding cassette (ABC) transporter that is required for cuticular lipid secretion. *Plant J*, 52: 485~498
- Borowski E, Bontemps-Gracz MM, Piwkowska A (2005). Strategies for overcoming ABC-transporters-mediated multidrug resistance (MDR) of tumor cells. *Acta Biochim Polonica*, 52: 609~627
- Chen S, Sánchez-Fernández R, Lyver ER, Dancis A, Rea PA (2007). Functional characterization of *AtATM1*, *AtATM2* and *AtATM3*, a subfamily of *Arabidopsis* half-molecule ATP-binding cassette transporters implicated in iron homeostasis. *J Biologic Chem*, 282: 21561~21571
- Choi H, Jin JY, Choi S, Hwang JU, Kim YY, Suh MC, Lee Y (2011). An ABCG/WBC-type ABC transporter is essential for transport of sporopollenin precursors for exine formation in developing pollen. *Plant J*, 65: 181~193
- Dudler R, Hertig C (1992). Structure of an *mdr*-like Gene from *Arabidopsis thaliana*. *J Biologic Chem*, 267: 5882~5888
- Ermolaeva MD, Wu M, Eisen JA, Salzberg SL (2003). The age of the *Arabidopsis thaliana* genome duplication. *Plant Mol Biol*, 51: 859~866
- Frelet-Barrand A, Kolukisaoglu HU, Plaza S, Ruffer M, Azevedo L, Hortensteiner S, Marinova K, Weder B, Schulz B, Klein M (2008). Comparative mutant analysis of *Arabidopsis* ABC-type ABC transporters: AtMRP2 contributes to detoxification, vacuolar organic anion transport and chlorophyll degradation. *Plant Cell Physiol*, 49: 557~569
- Heredia-Guerrero JA, San-Miguel MA, Sansom MS, Heredia A, Benitez JJ (2009). Chemical reactions in 2D: self-assembly and self-esterification of 9(10),16-dihydroxypalmitic acid on mica surface. *Langmuir*, 25: 6869~6874
- Janvilisri T, Venter H, Shahi S, Reuter G, Balakrishnan L, van Veen HW (2003). Sterol transport by the human breast cancer resistance protein (ABCG2) expressed in *Lactococcus lactis*. *J Biologic Chem*, 278: 20645~20651
- Kang J, Hwang JU, Lee M, Kim YY, Assmann SM, Martinoia E, Lee Y (2010). PDR-type ABC transporter mediates cellular uptake of the phytohormone abscisic acid. *Proc Natl Acad Sci USA*, 107: 2355~2360
- Klucken J, Buchler C, Orso E, Kaminski WE, Porsch-Ozcurumez M, Liebisch C, Kapinsky M, Diederich W, Drobnik W, Dean M et al (2000). *ABCG1* (*ABC8*), the human homolog of the *Drosophila* white gene, is a regulator of macrophage cholesterol and phospholipid transport. *Proc Natl Acad Sci USA*, 97: 817~822
- Kolattukudy PE (2001). Polyesters in higher plants. *Adv Biochem Biotechnol*, 71: 1~49
- Kuromori T, Miyaji T, Yabuuchi H, Shimizu H, Sugimoto E, Kamiya A, Moriyama Y, Shinozaki K (2010). ABC transporter AtABCG25 is involved in abscisic acid transport and responses. *Proc Natl Acad Sci USA*, 107: 2361~2366
- Lee M, Choi Y, Burla B, Kim YY, Jeon B, Maeshima M, Yoo JY, Martinoia E, Lee Y (2008). The ABC transporter AtABCB14 is a malate importer and modulates stomatal response to CO<sub>2</sub>. *Nat Cell Biol*, 10: 1217~1223
- Lee M, Lee K, Lee J, Noh EW, Lee Y (2005). *AtPDR12* contributes to lead resistance in *Arabidopsis*. *Plant Physiol*, 138: 827~836
- Linton KJ, Higgins CF (1998). The *Escherichia coli* ATP-binding cassette (ABC) proteins. *Mol Microbiol*, 28: 5~13
- Martinoia E, Klein M, Geisler M, Bovet L, Forestier C, Kolukisaoglu U, Muller-Rober B, Schulz B (2002). Multifunctionality of plant ABC transporters--more than just detoxifiers. *Planta*, 214: 345~355
- McFarlane HE, Shin JJ, Bird DA, Samuels AL (2010). *Arabidopsis* ABCG transporters, which are required for export of diverse cuticular lipids, dimerize in different combinations. *Plant Cell*, 22: 3066~3075
- Mentewab A, Stewart CN (2005). Overexpression of an *Arabidopsis thaliana* ABC transporter confers kanamycin resistance to transgenic plants. *Nat Biotechnol*, 23: 1177~1180
- Otsu CT, daSilva I, de Molfetta JB, da Silva LR, de Almeida-Engler J, Engler G, Torraca PC, Goldman GH, Goldman MH (2004).



- NtWBC1, an ABC transporter gene specifically expressed in tobacco reproductive organs. *J Exp Bot*, 55: 1643~1654
- Panikashvili D, Savaldi-Goldstein S, Mandel T, Yifhar T, Franke RB, Hofer R, Schreiber L, Chory J, Aharoni A (2007). The *Arabidopsis* DESPERADO/AtWBC11 transporter is required for cutin and wax secretion. *Plant Physiol*, 145: 1345~1360
- Pighin JA, Zheng H, Balakshin LJ, Goodman IP, Western TL, Jetter R, Kunst L, Samuels AL (2004). Plant cuticular lipid export requires an ABC transporter. *Science*, 306: 702~704
- Quilichini TD, Friedmann MC, Samuels AL, Douglas CJ (2010). ATP-binding cassette transporter G26 is required for male fertility and pollen exine formation in *Arabidopsis*. *Plant Physiol*, 154: 678~690
- Rea PA (2007). Plant ATP-binding cassette transporters. *Ann Rev Plant Biol*, 58: 347~375
- Sanchez-Fernandez R, Davies TG, Coleman JO, Rea PA (2001). The *Arabidopsis thaliana* ABC protein superfamily, a complete inventory. *J Genet Genom*, 276: 30231~30244
- Shitan N, Bazin I, Dan K, Obata K, Kigawa K, Ueda K, Sato F, Forestier C, Yazaki K (2003). Involvement of CjMDR1, a plant multidrug-resistance-type ATP-binding cassette protein, in alkaloid transport in *Coptis japonica*. *Proc Natl Acad Sci USA*, 100: 751~756
- Stukkens Y, Bultreys A, Grec S, Trombik T, Vanham D, Boutry M (2005). NpPDR1, a pleiotropic drug resistance-type ATP-binding cassette transporter from *Nicotiana plumbaginifolia*, plays a major role in plant pathogen defense. *Plant Physiol*, 139: 341~352
- Theodoulou FL (2000). Plant ABC transporters. *Bio Biophys Acta*, 1465: 79~103
- Titapiwatanakun B, Murphy AS (2009). Post-transcriptional regulation of auxin transport proteins: cellular trafficking, protein phosphorylation, protein maturation, ubiquitination, and membrane composition. *J Exp Bot*, 60: 1093~1107
- Ukitsu H, Kuromori T, Toyooka K, Goto Y, Matsuoka K, Sakuradani E, Shimizu S, Kamiya A, Imura Y, Yuguchi M et al (2007). Cytological and biochemical analysis of *COF1*, an *Arabidopsis* mutant of an ABC transporter gene. *Plant Cell Physiol*, 48: 1524~1533
- Verrier PJ, Bird D, Burla B, Dassa E, Forestier C, Geisler M, Klein M, Kolukisaoglu U, Lee Y, Martinoia E et al (2008). Plant ABC proteins—a unified nomenclature and updated inventory. *Trends Plant Sci*, 13: 151~159
- Wang C, Marshall A, Zhang D, Wilson ZA (2012). ANAP: an integrated knowledge base for *Arabidopsis* protein interaction network analysis. *Plant Physiol*, 158: 1523~1533
- Xu J, Yang CS, Yuan Z, Zhang DB, Gondwe MY, Ding ZW, Liang WQ, Wilson ZA (2010). The *ABORTED MICROSPORES* regulatory network is required for postmeiotic male reproductive development in *Arabidopsis thaliana*. *Plant Cell*, 22: 91~107
- Xu X, Qiu J, Xu Y, Xu CW (2012). Molecular evolution and expression analysis of subfamily ABCB transporter genes in rice. *Chinese J Rice Sci*, 26: 127~136
- Xu XH, Zhao HJ, Liu QL, Frank T, Engel KH, An G, Shu QY (2009). Mutations of the multi-drug resistance-associated protein ABC transporter gene 5 result in reduction of phytic acid in rice seeds. *Theor Appl Genet*, 119: 75~83
- Zhang DB, Luo X, Zhu L (2011). Cytological analysis and genetic control of rice anther development. *J Genet Genom*, 38: 379~390
- Zhu YQ, Xu KX, Luo B, Wang JW, Chen XY (2003). An ATP-binding cassette transporter GhWBC1 from elongating cotton fibers. *Plant Physiol*, 133: 580~588